

09/582328

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
 COMPLIANCE WITH RULE 17.1(a) OR (b)



REC'D	24 FEB 1999
WIPO	PCT

Bescheinigung

ESO

Herr Professor Dr. Wolf-Georg F o r s s m a n n in Hannover/Deutschland hat
 eine Patentanmeldung unter der Bezeichnung

"Serin-Proteinase-Inhibitoren"

am 8. Januar 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprüng-
 lichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole
 C 07 K, A 61 K und C 12 N der Internationalen Patentklassifikation erhalten.

München, den 15. Januar 1999

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Faust

Aktenzeichen: 198 00 363.3

Faust

11 04.02.99

08. Januar 1998

980114de Me-Sch/kk

Serin-Proteinase-Inhibitoren

Die Erfindung betrifft Serin-Proteinase-Inhibitoren, cDNA kodierend für Serin-Proteinase-Inhibitoren, Arzneimittel enthaltend die Inhibitoren oder deren codierende Nucleinsäure, Verwendungen der erfindungsgemäßen Verbindungen zur Herstellung von Arzneimitteln zur Behandlung verschiedener Indikationen, Antikörper oder Antikörperfragmente gegen Epitope der erfindungsgemäßen Verbindungen, Poly- oder Oligonucleotide, die mit Genen der erfindungsgemäßen Verbindungen hybridisieren, ein Diagnostikum zum Aufspüren der erfindungsgemäßen Verbindungen, sowie Arzneimittel enthaltend Antikörper oder Poly- oder Oligonucleotide gemäß der Erfindung.

Proteolytische Prozesse spielen in allen Organismen eine bedeutende physiologische Rolle, wobei zwischen unspezifischen und spezifischen proteolytischen Reaktionen zu unterscheiden ist. Zu den ersten gehören beispielsweise der Nahrungsaufschluß im Verdauungstrakt durch Endopeptidasen sowie der intrazelluläre Abbau verbrauchter endogener Substanzen und phagozytierten Materials durch lysosomale Proteinasen. Spezifische Proteolysen dienen meistens der Überführung eines Proenzyms in die aktive Form wie bei der

Überführung von Trypsinogen in Trypsin und Chymotrypsinogen in Chymotrypsin sowie bei den Kallikrein-Kinin-Kaskaden und der Blutgerinnungskaskade. Je nach Beschaffenheit des reaktiven Zentrums der daran beteiligten Proteinasen werden diese in die Klassen der Serin-Proteinasen (z.B. Chymotrypsin, Trypsin, Elastase und Kathepsin G), der Aspartat-Proteinasen (z.B. Kathepsin D, Kathepsin E und Pepsin), der Cystein-Proteinasen (z.B. Kathepsin B, Kathepsin H und Kathepsin L) und der Metallo-Proteinasen (z.B. Kollagenase und Thermolysin) unterteilt.

Um die oft kaskadenartig verlaufenden proteolytischen Prozesse gegenregulieren zu können, verfügt der Organismus über eine Reihe von anderen Proteinen, den Proteinase-Inhibitoren (zur Übersicht siehe Laskowski und Kato, 1980 und Bode und Huber, 1992). So schützen die in der Leber synthetisierten, humanen Plasma-Proteinase-Inhibitoren α_1 -Antichymotrypsin und α_1 -Proteinase-Inhibitoren das Lungengewebe vor unspezifischem Angriff durch die Proteinasen Kathepsin G bzw. Elastase aus polymorphkernigen Lymphozyten. Bei einem Ungleichgewicht zwischen Proteinasen und ihren spezifischen Inhibitoren kann es zum Auftreten pathologischer Effekte kommen. Ein übermäßiges Verhältnis von Elastase zu α_1 -Proteinase-Inhibitor erhöht beispielsweise bei Patienten mit genetisch bedingtem Mangel an diesem Faktor das Risiko der Bildung eines Lungenemphysems um ca. 20 bis 30fach gegenüber der Normalbevölkerung (Carrel und Owen, 1980). Bei Rauchern wird die Emphysembildung mittels Oxidation der im reaktiven Zentrum des α_1 -Proteinase-Inhibitors befindlichen Aminosäure Methionin durch im Zigarettenrauch enthaltene Oxidantien begünstigt (Miller und Kushner, 1969; Ohlsson et al., 1980). Auch im Falle der Infektion mit Gram-negativen Bakterien können deren Endotoxine eine Desintegration von Phagozyten und damit die Ausschüttung lysosomaler Proteinasen verursachen, was durch den erhöhten Verbrauch an Proteinase-Inhibitoren unkontrollierte Gewebsschädigung und Entzündungen verursachen kann. Aus diesem Grund besitzen bestimmte

Proteinase-Inhibitoren ein hohes therapeutisches Potential (siehe z.B. Fritz, 1980).

Das der Erfindung zugrunde liegende technische Problem besteht darin, weitere Inhibitoren von Serin-Proteinasen zur Verfügung zu stellen. Des weiteren sollten die für die erfindungsgemäßen Inhibitoren kodierenden Gene bzw. cDNA zur Verfügung gestellt werden.

Spezifisches Merkmal der erfindungsgemäßen Serin-Proteinase-Inhibitoren ist, daß der Serin-Proteinase-Inhibitor eine Domäne mit vier Cysteinen aufweist und sich zwischen dem ersten und einem zweiten Cystein eine Sequenz von 0 bis 20 Aminosäuren befindet oder der Serin-Proteinase-Inhibitor eine Domäne mit sechs Cysteinen aufweist und sich zwischen dem ersten und zweiten Cystein eine Sequenz von 7 bis 20 Aminosäuren befindet.

Bevorzugterweise befindet sich zwischen einem ersten und einem zweiten Cystein eine Sequenz von 13 Aminosäuren und/oder zwischen einem zweiten und einem dritten Cystein eine Sequenz von 18 Aminosäuren und/oder zwischen einem dritten und vierten Cystein eine Sequenz von 2 Aminosäuren.

Es wird insbesondere bevorzugt, daß die Sequenz zwischen einem ersten und zweiten Cystein ausgewählt wird aus

HEFQAFMKNGKLF,
DDFKKGERDGDFI,
SAFRPFVRNGRLG,
KEYEKQVRNGRLF,
SQYQNQAKNGILF,
NEYRKLVRNGKLA,

SEYRKSRRKNGRLF,
SEFRDQVRNGTLI,
SEYRHVVRNGRLP,
DEFRRLLQNGKLF,
AEYREQMKNGRLS oder
DEFRSQMKNGLI

und/oder die Sequenz zwischen einem zweiten und dritten Cystein ausgewählt wird aus

PQDKKFFQSLDGIMFINK,	TRENDPIQGPDGKMHGNT,
TRENDPVLGPDGKTHGNK,	TREHNPVRGPDGKMHGNK,
TRESDPVRGPDGRMHGNK,	TRENDPIEGLDGKIHGNT,
TRENDPIRGPDGKMHGNL,	TRENDPVRGPDGKTHGNK,
TRENDPIQGPDGKVHGNT,	TRESDPVRDADGKSYNNQ oder
	TRESDPVRGPDGKTHGNK

und/oder die Sequenz zwischen einem dritten und vierten Cystein ausgewählt wird aus

AT, AL, AM, SM oder TM.

Besonders bevorzugt wird, daß der erfindungsgemäße Serin-Proteinase-Inhibitor einer der Formeln

R_1 -C-HEFQAFMKNGKLF-C-PQDKKFFQSLDGIMFINK-C-AT-C- R_2
 R_1 -C-DDFKKGERDGDFI-C-PDYIEAVCGTDGKTYDNR-C-AL-C- R_2
 R_1 -C-SAFRPFVRNGRLG-C-TRENDPVLGPDGKTHGNK-C-AM-C- R_2
 R_1 -C-KEYEKQVRNGRLF-C-TRESDPVRGPDGRMHGNK-C-AL-C- R_2
 R_1 -C-SQYQNQAKNGILF-C-TRENDPIRGPDGKMHGNL-C-SM-C- R_2
 R_1 -C-NEYRKLVRNGKLA-C-TRENDPIQGPDGKVHGNT-C-SM-C- R_2
 R_1 -C-SEYRKSRRKNGRLF-C-TRENDPIQGPDGKMHGNT-C-SM-C- R_2
 R_1 -C-SEFRDQVRNGTLI-C-TREHNPVRGPDGKMHGNK-C-AM-C- R_2
 R_1 -C-SEYRHYVRNGRLP-C-TRENDPIEGLDGKIHGNT-C-SM-C- R_2
 R_1 -C-DEFRRLLQNGKLF-C-TRENDPVRGPDGKTHGNK-C-AM-C- R_2
 R_1 -C-AEYREQMKNGRLS-C-TRESDPVRDADGKSYNNQ-C-TM-C- R_2
 R_1 -C-DEFRSQMKNGKLI-C-TRESDPVRGPDGKTHGNK-C-TM-C- R_2 ,

worin R_1 NH_2 , eine Aminosäure oder ein Peptid mit bis zu 100 Aminosäuren ist und R_2 $COOH$, $CONH_2$, eine Aminosäure oder ein Peptid mit bis zu 100 Aminosäuren ist, entspricht.

Es ist weiterhin bevorzugt, daß die Serin-Proteinase ein oder mehrere Disulfidbrücken aufweist. Dabei ist besonders bevorzugt, daß er zwischen dem ersten und vierten Cystein und/oder dem zweiten und dritten Cystein eine Disulfidbrücke aufweist oder daß er zwischen dem ersten und fünften Cystein und/oder

dem zweiten und vierten Cystein und/oder dem dritten und sechsten Cystein eine Disulfidbrücke aufweist.

Bevorzugte Vertreter der erfindungsgemäßen Serin-Proteinase-Inhibitoren sind die Verbindungen HF 6479 und HF 7665 sowie Fragmente der Proteine VAKTI-1 und VAKTI-2 gemäß den Figuren 1 und 2.

Aus den Figuren 1 bis 3 lassen sich neben der Aminosäuresequenz der erfindungsgemäß bevorzugten Verbindungen auch weitere Informationen bezüglich der cDNA, die für die erfindungsgemäßen Verbindungen kodiert, entnehmen. Insbesondere werden die entsprechenden Motive und Primer hybridisierenden Stellen angegeben.

Die erfindungsgemäße Verbindung HF 3479 weist eine Masse von 6.479 Dalton auf, diejenige von HF 7665 beträgt 7.665 Dalton, beide wurden aus Hämofiltrat aufgereinigt.

Erfindungsgemäß beansprucht wird auch eine cDNA, kodierend für die erfindungsgemäßen Verbindungen, insbesondere eine cDNA mit der Nucleinsäuresequenz gemäß Figuren 1 bis 2.

Die erfindungsgemäßen Verbindungen sind als Arzneimittel geeignet. Gegebenenfalls werden sie zusammen mit pharmazeutisch verträglichen Trägerstoffen appliziert.

Die erfindungsgemäßen Arzneimittel enthaltend die erfindungsgemäßen Proteinase-Inhibitoren werden vorzugsweise in Mengen von 1 bis 100 mg/kg Körpergewicht des Patienten verabreicht. Als Verabreichungsform kommen alle galenischen Zubereitungen für Peptidwirkstoffe in Frage. Die Arzneimittel enthaltend Nucleinsäuren gemäß der Erfindung werden vorzugsweise in Mengen von 0,1 bis 100 mg/kg Körpergewicht eines entsprechenden Patienten verabreicht. Hier kommen als galenische Verabreichungsformen solche in Betracht, die zur Applikation von Nucleinsäuren geeignet sind, ohne daß die Nucleinsäuren

vor Erreichen des Wirkortes durch Stoffwechseleinflüsse unwirksam gemacht werden. Als galenische Verabreichungsform können z.B. Liposomen eingesetzt werden, in denen die Nucleinsäuren befindlich sind.

Die erfindungsgemäßen Verbindungen kommen insbesondere zur Behandlung von akuten oder chronischen Cervixentzündungen, Entzündungen der Bartholinschen Drüsen oder anderer vaginaler Bereiche, Tonsillitis, Pharyngitis und Laryngitis, mit exzessiver Schleimbildung verbundener akut oder chronisch entzündlicher Prozesse und sich daraus ergebender akuter Not-situationen, postoperativer Blutungen aufgrund Hyperfibrinolyse sowie zur Prophylaxe der Lungenemphysembildung bei α_1 -Proteinase-Inhibitor-mangel in Betracht.

Die erfindungsgemäßen Verbindungen können bei Mangel an Serin-Proteinase-Inhibitoren verabreicht werden, um endogene Defizite auszugleichen. Die Nucleinsäuren können, direkt oder an geeignete Vehikel gekoppelt, auch zum Einsatz in der Gentherapie gelangen. Als geeignete Vektoren kommen insbesondere attenuierte Adenoviren, in die entsprechenden Gene inkorporiert werden, in Frage.

Die erfindungsgemäßen Polypeptide, insbesondere VAKTI-I und VAKTI-II, können zur Herstellung von Antikörpern oder Antikörperfragmenten dienen. Diese werden in einfacher Weise durch Immunisierung geeigneter Säuger hergestellt. Durch an sich bekannte Operationen können die Antikörper auch humanisiert werden, so daß diese Antikörper ebenfalls zum therapeutischen Einsatz gelangen können. Antikörper oder Antikörperfragmente können dann zur Regulation von Erkrankungen eingesetzt werden, bei denen die Proteinase-Inhibitoren pathologisch exprimiert werden. Ebenso können zu den erfindungsgemäßen Nucleinsäuren komplementäre Antisense-Nucleinsäuren zum therapeutischen Einsatz bei Überexpression der Proteinase-Inhibitor-Genen eingesetzt werden.

Die erfindungsgemäßen Verbindungen sind in einfacher Weise durch an sich bekannte Methoden der Peptid- bzw. Nucleotidsynthese herstellbar. Einer gentechnischen Herstellung der Verbindungen steht ebenfalls nichts im Wege.

Der Fachmann erkennt, daß bei den Polypeptiden gemäß der Erfindung auch Fragmente verwendet werden können, sofern sie die inhibitorischen Eigenschaften der Serin-Proteinase-Inhibitoren beibehalten. Das Auffinden solcher Fragmente ist dem Fachmann bekannt. So erfolgt dies beispielsweise durch gezielte enzymatische Spaltung der erfindungsgemäßen Verbindungen. Es können auch in den Seitenketten modifizierte Aminosäuren eingesetzt werden. Auch N- und C-terminal modifizierte Polypeptide kommen in Betracht. Insbesondere können phosphorylierte, glycosylierte, methylierte, acetylierte oder in ähnlicher Weise modifizierte Polypeptide eingesetzt werden, sofern sie die Wirkung der Serin-Proteinase-Inhibitoren nicht relevant beeinträchtigen.

Bei den Nucleinsäuren gemäß der Erfindung kommen auch Derivate in Betracht, die je nach Codon Usage modifizierte Tripletstrukturen aufweisen. Desweiteren sind als Nucleinsäuren gemäß der Erfindung auch solche zu verstehen, die durch Nucleasen gegenüber den nativen Verbindungen weniger stark abgebaut werden, beispielsweise die entsprechenden SODN-Derivate, die in der Antisense-Technologie üblicherweise eingesetzt werden, um die Antisense-Strukturen gegenüber enzymatischen Angriffen stabiler auszugestalten.

Auch mit den Polypeptiden homologe Strukturen kommen in Betracht. Dies sind insbesondere Polypeptidstrukturen, bei denen Aminosäuren ausgetauscht sind. So können beispielsweise konservative Aminosäuresubstitutionen in hochkonservierten Regionen wie folgt berücksichtigt werden: Jede Isoleucin-, Valin- und Leucin-Aminosäure kann gegen eine andere dieser Aminosäuren ausgetauscht sein, Aspartat kann gegen Glutamat und umgekehrt ausgetauscht sein, Glutamin gegen Asparagin

und umgekehrt, Serin gegen Trionin und umgekehrt. Konservative Aminosäuresubstitutionen in weniger hochkonservierten Regionen können wie folgt sein: Jede der Aminosäuren Isoleucin, Valin und Leucin gegen jede andere Aminosäuren, Aspartat gegen Glutamat und umgekehrt, Glutamin gegen Asparagin und umgekehrt, Serin gegen Theonin und umgekehrt, Glycin gegen Alanin und umgekehrt, Alanin gegen Valin und umgekehrt, Methionin gegen jede der Aminosäuren Leucin, Isoleucin oder Valin, Lysin gegen Arginin und umgekehrt, eine der Aminosäuren Aspartat oder Glutamat gegen eine der Aminosäuren Arginin oder Lysin, Histidin gegen eine der Aminosäuren Arginin oder Lysin, Glutamin gegen Glutamat und umgekehrt und Asparagin gegen Aspartat und umgekehrt.

Die Wirkungsweise der erfindungsgemäßen Peptide wird durch das folgende Beispiel erläutert.

Beispiel

Messung der Proteinase Inhibition durch HF 7665

Meßansatz:

84 µl	Meßpuffer (0,1 M HEPES, pH 7.5; 0,5 M NaCl ₂)
1 µl	Trypsin (1 mg/ml in 1 mM HCl, 20 mM CaCl ₂)
5 µl	L-BABNA (6 mg/ml N α -Benzoyl-L-Arginine-p-Nitro-anilide Hydrochloride)
10 µl	Proteinase-Inhibitor (10 µM bzw. 75 µg/ml HF 7665 in H ₂ O).

Die Reaktion wurde durch Zugabe des chromogenen Substrates gestartet und der Substratumsatz mittels Photometer bei $\lambda = 405$ nm verfolgt. Nach ca. fünf Minuten wurden 10 µl Proteinase-Inhibitor bzw. entsprechende Kontrollen dazugegeben und der weitere Extinktionsverlauf beobachtet.

Es konnte gezeigt werden, daß HF 7665 in einer Endkonzentration von ca. $1 \mu\text{M}$ bzw. $7,5 \mu\text{g/ml}$ einen inhibitorischen Effekt auf Trypsin besitzt. Kontrollversuche mit entsprechenden Mengen an BSA ($7,5 \mu\text{g/ml}$) und Acetonitril/TFA (0,8% ACN/0,001% TFA) zeigten keine Trypsininhibierung. Weiterhin konnte kein inhibitorischer Effekt von HF 7665 auf Chymotrypsin bei einem ähnlichen Test beobachtet werden.

Figur 3 zeigt, daß sich nach Zugabe von HF 7665 der Substratumsatz durch Trypsininhibierung um ca. 30% vermindert.

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

- (A) NAME: Prof. Dr. Wolf-Georg Forssmann
- (B) STRASSE: Feodor-Lynen-Str. 31
- (C) ORT: Hannover
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 30625

(ii) BEZEICHNUNG DER ERFINDUNG: Serin-Proteinase-Inhibitoren

(iii) ANZAHL DER SEQUENZEN: 34

(iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 177 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

Met	Lys	Ile	Ala	Thr	Val	Ser	Val	Leu	Leu	Pro	Leu	Ala	Leu	Cys	Leu	1	5	10	15
Ile	Gln	Asp	Ala	Ala	Ser	Lys	Asn	Glu	Asp	Gln	Glu	Met	Cys	His	Glu	20	25	30	
Phe	Gln	Ala	Phe	Met	Lys	Asn	Gly	Lys	Leu	Phe	Cys	Pro	Gln	Asp	Lys	35	40	45	
Lys	Phe	Phe	Gln	Ser	Leu	Asp	Gly	Ile	Met	Phe	Ile	Asn	Lys	Cys	Ala	50	55	60	
Thr	Cys	Lys	Met	Ile	Leu	Glu	Lys	Glu	Ala	Lys	Ser	Gln	Lys	Arg	Ala	65	70	75	80
Arg	His	Leu	Ala	Arg	Ala	Pro	Lys	Ala	Thr	Ala	Pro	Thr	Glu	Leu	Asn	85	90	95	
Cys	Asp	Asp	Phe	Lys	Lys	Gly	Glu	Arg	Asp	Gly	Asp	Phe	Ile	Cys	Pro	100	105	110	
Asp	Tyr	Tyr	Glu	Ala	Val	Cys	Gly	Thr	Asp	Gly	Lys	Thr	Tyr	Asp	Asn	115	120	125	
Arg	Cys	Ala	Leu	Cys	Ala	Glu	Asn	Ala	Lys	Thr	Gly	Ser	Gln	Ile	Gly	130	135	140	
Val	Lys	Ser	Glu	Gly	Glu	Cys	Lys	Ser	Ser	Asn	Pro	Glu	Gln	Val	Arg	145	150	155	160

11 04.02.99

- 11 -

Ser Ile Val Ser Leu Met Gly Asn Thr Gly Arg Leu Thr Ser Asn Ser
165 170 175

Lys

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 922 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met	Lys	Ile	Ala	Thr	Val	Ser	Val	Leu	Leu	Pro	Leu	Ala	Leu	Cys	Leu	1	5	10	15
Ile	Gln	Asp	Ala	Ala	Ser	Lys	Asn	Glu	Asp	Gln	Glu	Met	Cys	His	Glu	20	25	30	
Phe	Gln	Ala	Phe	Met	Lys	Asn	Gly	Lys	Leu	Phe	Cys	Pro	Gln	Asp	Lys	35	40	45	
Lys	Phe	Phe	Gln	Ser	Leu	Asp	Gly	Ile	Met	Phe	Ile	Asn	Lys	Cys	Ala	50	55	60	
Thr	Cys	Lys	Met	Ile	Leu	Glu	Lys	Glu	Ala	Lys	Ser	Gln	Lys	Arg	Ala	65	70	75	80
Arg	His	Leu	Ala	Arg	Ala	Pro	Lys	Ala	Thr	Ala	Pro	Thr	Glu	Leu	Asn	85	90	95	
Cys	Asp	Asp	Phe	Lys	Lys	Gly	Glu	Arg	Asp	Gly	Asp	Phe	Ile	Cys	Pro	100	105	110	
Asp	Tyr	Tyr	Glu	Ala	Val	Cys	Gly	Thr	Asp	Gly	Lys	Thr	Tyr	Asp	Asn	115	120	125	
Arg	Cys	Ala	Leu	Cys	Ala	Glu	Asn	Ala	Lys	Thr	Gly	Ser	Gln	Ile	Gly	130	135	140	
Val	Lys	Ser	Glu	Gly	Glu	Cys	Lys	Ser	Ser	Asn	Pro	Glu	Gln	Asp	Val	145	150	155	160
Cys	Ser	Ala	Phe	Arg	Pro	Phe	Val	Arg	Asn	Gly	Arg	Leu	Gly	Cys	Thr	165	170	175	
Arg	Glu	Asn	Asp	Pro	Val	Leu	Gly	Pro	Asp	Gly	Lys	Thr	His	Gly	Asn	180	185	190	
Lys	Cys	Ala	Met	Cys	Ala	Glu	Leu	Phe	Leu	Lys	Glu	Ala	Glu	Asn	Ala	195	200	205	
Lys	Arg	Glu	Gly	Glu	Thr	Arg	Ile	Arg	Arg	Asn	Ala	Glu	Lys	Asp	Phe	210	215	220	
Cys	Lys	Glu	Tyr	Glu	Lys	Gln	Val	Arg	Asn	Gly	Arg	Leu	Phe	Cys	Thr	225	230	235	240

01.02.99

- 12 -

Arg	Glu	Ser	Asp	Pro	Val	Arg	Gly	Pro	Asp	Gly	Arg	Met	His	Gly	Asn
				245					250					255	
Lys	Cys	Ala	Leu	Cys	Ala	Glu	Ile	Phe	Lys	Arg	Arg	Phe	Ser	Glu	Glu
			260					265					270		
Asn	Ser	Lys	Thr	Asp	Gln	Asn	Leu	Gly	Lys	Ala	Glu	Glu	Lys	Thr	Lys
		275					280					285			
Val	Lys	Arg	Glu	Ile	Val	Lys	Leu	Cys	Ser	Gln	Tyr	Gln	Asn	Gln	Ala
	290					295					300				
Lys	Asn	Gly	Ile	Leu	Phe	Cys	Thr	Arg	Glu	Asn	Asp	Pro	Ile	Arg	Gly
305					310					315					320
Pro	Asp	Gly	Lys	Met	His	Gly	Asn	Leu	Cys	Ser	Met	Cys	Gln	Val	Tyr
				325					330					335	
Phe	Gln	Ala	Glu	Asn	Glu	Glu	Lys	Lys	Lys	Ala	Glu	Ala	Arg	Ala	Arg
			340					345					350		
Asn	Lys	Arg	Glu	Ser	Gly	Lys	Ala	Thr	Ser	Tyr	Ala	Glu	Leu	Cys	Asn
		355					360					365			
Glu	Tyr	Arg	Lys	Leu	Val	Arg	Asn	Gly	Lys	Leu	Ala	Cys	Thr	Arg	Glu
	370					375					380				
Asn	Asp	Pro	Ile	Gln	Gly	Pro	Asp	Gly	Lys	Val	His	Gly	Asn	Thr	Cys
385					390					395					400
Ser	Met	Cys	Glu	Val	Phe	Phe	Gln	Ala	Glu	Glu	Glu	Glu	Lys	Lys	Lys
				405					410					415	
Lys	Glu	Gly	Glu	Ser	Arg	Asn	Lys	Arg	Gln	Ser	Lys	Ser	Thr	Ala	Ser
			420					425					430		
Phe	Glu	Glu	Leu	Cys	Ser	Glu	Tyr	Arg	Lys	Ser	Arg	Lys	Asn	Gly	Arg
		435					440					445			
Leu	Phe	Cys	Thr	Arg	Glu	Asn	Asp	Pro	Ile	Gln	Gly	Pro	Asp	Gly	Lys
	450					455					460				
Met	His	Gly	Asn	Thr	Cys	Ser	Met	Cys	Glu	Ala	Phe	Phe	Gln	Gln	Glu
465					470					475					480
Glu	Arg	Ala	Arg	Ala	Lys	Ala	Lys	Arg	Glu	Ala	Ala	Lys	Glu	Ile	Cys
				485					490					495	
Ser	Glu	Phe	Arg	Asp	Gln	Val	Arg	Asn	Gly	Thr	Leu	Ile	Cys	Thr	Arg
			500					505					510		
Glu	His	Asn	Pro	Val	Arg	Gly	Pro	Asp	Gly	Lys	Met	His	Gly	Asn	Lys
		515					520					525			
Cys	Ala	Met	Cys	Ala	Ser	Val	Phe	Lys	Leu	Glu	Glu	Glu	Glu	Lys	Lys
	530					535					540				
Asn	Asp	Lys	Glu	Glu	Lys	Gly	Lys	Val	Glu	Ala	Glu	Lys	Val	Lys	Arg
545					550					555					560
Glu	Ala	Val	Gln	Glu	Leu	Cys	Ser	Glu	Tyr	Arg	His	Tyr	Val	Arg	Asn
			565						570					575	
Gly	Arg	Leu	Pro	Cys	Thr	Arg	Glu	Asn	Asp	Pro	Ile	Glu	Gly	Leu	Asp
			580					585					590		

Gly	Lys	Ile	His	Gly	Asn	Thr	Cys	Ser	Met	Cys	Glu	Ala	Phe	Phe	Gln	595	600	605
Gln	Glu	Ala	Lys	Glu	Lys	Glu	Arg	Ala	Glu	Pro	Arg	Ala	Lys	Val	Lys	610	615	620
Arg	Glu	Ala	Glu	Lys	Glu	Thr	Cys	Asp	Glu	Phe	Arg	Arg	Leu	Leu	Gln	625	630	635
Asn	Gly	Lys	Leu	Phe	Cys	Thr	Arg	Glu	Asn	Asp	Pro	Val	Arg	Gly	Pro	645	650	655
Asp	Gly	Lys	Thr	His	Gly	Asn	Lys	Cys	Ala	Met	Cys	Lys	Ala	Val	Phe	660	665	670
Gln	Lys	Glu	Asn	Glu	Glu	Arg	Lys	Arg	Lys	Glu	Glu	Glu	Asp	Gln	Arg	675	680	685
Asn	Ala	Ala	Gly	His	Gly	Ser	Ser	Gly	Gly	Gly	Gly	Gly	Asn	Thr	Gln	690	695	700
Asp	Glu	Cys	Ala	Glu	Tyr	Arg	Glu	Gln	Met	Lys	Asn	Gly	Arg	Leu	Ser	705	710	715
Cys	Thr	Arg	Glu	Ser	Asp	Pro	Val	Arg	Asp	Ala	Asp	Gly	Lys	Ser	Tyr	725	730	735
Asn	Asn	Gln	Cys	Thr	Met	Cys	Lys	Ala	Lys	Leu	Glu	Arg	Glu	Ala	Glu	740	745	750
Arg	Lys	Asn	Glu	Tyr	Ser	Arg	Ser	Arg	Ser	Asn	Gly	Thr	Gly	Ser	Glu	755	760	765
Ser	Gly	Lys	Asp	Thr	Cys	Asp	Glu	Phe	Arg	Ser	Gln	Met	Lys	Asn	Gly	770	775	780
Lys	Leu	Ile	Cys	Thr	Arg	Glu	Ser	Asp	Pro	Val	Arg	Gly	Pro	Asp	Gly	785	790	795
Lys	Thr	His	Gly	Asn	Lys	Cys	Thr	Met	Cys	Lys	Glu	Lys	Leu	Glu	Arg	805	810	815
Glu	Ala	Ala	Glu	Lys	Lys	Arg	Lys	Arg	Met	Lys	Thr	Gly	Ala	Ile	Gln	820	825	830
Glu	Lys	Gly	Ala	Ile	Gln	Glu	Lys	Gly	Ala	Met	Thr	Lys	Arg	Ile	Cys	835	840	845
Val	Val	Asn	Phe	Glu	Ala	Cys	Arg	Glu	Met	Glu	Ser	Leu	Ser	Ala	Pro	850	855	860
Glu	Lys	Ile	Thr	Leu	Phe	Glu	Ala	His	Met	Ala	Arg	Cys	Thr	Ser	Ile	865	870	875
Asn	Val	Leu	Cys	Val	Arg	Ala	Ser	Leu	Ile	Glu	Lys	Leu	Met	Lys	Glu	885	890	895
Lys	Arg	Lys	Met	Lys	Arg	Asn	Gln	Val	Ala	Ser	Pro	Gln	Ile	Met	Gln	900	905	910
Arg	Met	Ser	Ala	Val	Asn	Phe	Glu	Thr	Ile							915	920	

(2) ANGABEN ZU SEQ ID NO: 3:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 55 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

Lys Asn Glu Asp Gln Glu Met Cys His Glu Phe Gln Ala Phe Met Lys
 1 5 10 15
 Asn Gly Lys Leu Phe Cys Pro Gln Asp Lys Lys Phe Phe Gln Ser Leu
 20 25 30
 Asp Gly Ile Met Phe Ile Asn Lys Cys Ala Thr Cys Lys Met Ile Leu
 35 40 45
 Glu Lys Glu Ala Lys Ser Gln
 50 55

(2) ANGABEN ZU SEQ ID NO: 4:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 68 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Glu Ser Gly Lys Ala Thr Ser Tyr Ala Glu Leu Cys Asn Glu Tyr Arg
 1 5 10 15
 Lys Leu Val Arg Asn Gly Lys Leu Ala Cys Thr Arg Glu Asn Asp Pro
 20 25 30
 Ile Gln Gly Pro Asp Gly Lys Val His Gly Asn Thr Cys Ser Met Cys
 35 40 45
 Glu Val Phe Phe Gln Ala Glu Glu Glu Glu Lys Lys Lys Lys Glu Gly
 50 55 60
 Glu Ser Arg Asn
 65

(2) ANGABEN ZU SEQ ID NO: 5:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 748 Basenpaare
 (B) ART: Nucleotid
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

ATGCATGGAG TGGACCTGTA GGCGACTTGC ATCGTCTTCA ACATGAAGAT AGCCACAGTG	60
TCAGTGCTTC TGCCCTTGGC TCTTTGCCTC ATACAAGATG CTGCCAGTAA GAATGAAGAT	120
CAGGAAATGT GCCATGAATT TCAGGCATTT ATGAAAAATG GAAAACTGTT CTGTCCCCAG	180
GATAAGAAAT TTTTTCAAAG TCTTGATGGA ATAATGTTCA TCAATAAATG TGCCACGTGC	240
AAAATGATAC TGGAAAAAGA AGCAAAATCA CAGAAGAGGG CCAGGCATTT AGCAAGAGCT	300
CCCAAGGCTA CTGCCCCAAC AGAGCTGAAT TGTGATGATT TTAAAAAAGG AGAAAGAGAT	360
GGGGATTTTA TCTGTCCTGA TTATTATGAA GCTGTTTGTG GCACAGATGG GAAAACATAT	420
GACAACAGAT GTGCACTGTG TGCTGAGAAT GCGAAAACCG GGTCCCAAAT TGGTGTAATA	480
AGTGAAGGGG AATGTAAGAG CAGTAATCCA GAGCAGGTGA GGTCAATTGT CAGCCTGATG	540
GGAAATACTG GGAGGCTAAC TTCAAATAGT AAGTAGGTGC TGTCTCTTC CTTCTTAGGT	600
GGGAGCCTTG GAAGGAATTA ATTCTTGCTT TATGTGAAAT GGAATACCCA GTTACTGCCC	660
ACTAATATGA AAAAGCTAAT TATAGTCTCT GAAACTGGAT CAGATTACTT TGGTGGTTAA	720
GATCTTTCAA TCTATTGCTG CTTTGTAT	748

(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 3531 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: cDNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

ATGCATGGAG TGGACCTGTA GGCGACTTGC ATCGTCTTCA ACATGAAGAT AGCCACAGTG	60
TCAGTGCTTC TGCCCTTGGC TCTTTGCCTC ATACAAGATG CTGCCAGTAA GAATGAAGAT	120
CAGGAAATGT GCCATGAATT TCAGGCATTT ATGAAAAATG GAAAACTGTT CTGTCCCCAG	180
GATAAGAAAT TTTTTCAAAG TCTTGATGGA ATAATGTTCA TCAATAAATG TGCCACGTGC	240
AAAATGATAC TGGAAAAAGA AGCAAAATCA CAGAAGAGGG CCAGGCATTT AGCAAGAGCT	300
CCCAAGGCTA CTGCCCCAAC AGAGCTGAAT TGTGATGATT TTAAAAAAGG AGAAAGAGAT	360
GGGGATTTTA TCTGTCCTGA TTATTATGAA GCTGTTTGTG GCACAGATGG GAAAACATAT	420
GACAACAGAT GTGCACTGTG TGCTGAGAAT GCGAAAACCG GGTCCCAAAT TGGTGTAATA	480
AGTGAAGGGG AATGTAAGAG CAGTAATCCA GAGCAGGATG TATGCAGTGC TTTTCGGCCC	540
TTTGTTAGAA ATGGAAGACT TGGATGCACA AGGGAAAATG ATCCTGTTCT TGGTCCTGAT	600
GGGAAGACGC ATGGCAATAA GTGTGCAATG TGTGCTGAGC TGTTTTTAAA AGAAGCTGAA	660



AATGCCAAGC	GAGAGGGTGA	AACTAGAATT	CGACGAAATG	CTGAAAAGGA	TTTTTGCAAG	720
GAATATGAAA	AACAAGTGAG	AAATGGAAGG	CTTTTTTGTA	CACGGGAGAG	TGATCCAGTC	780
CGTGGCCCTG	ACGGCAGGAT	GATGGCAAC	AAATGTGCCC	TGTGTGCTGA	AATTTTCAAG	840
CGGCGTTTTT	CAGAGGAAAA	CAGTAAAACA	GATCAAAATT	TGGGAAAAGC	TGAAGAAAAA	900
ACTAAAGTTA	AAAGAGAAAT	TGTGAAACTC	TGCAGTCAAT	ATCAAAATCA	GGCAAAGAAT	960
GGAATACTTT	TCTGTACCAG	AGAAAATGAC	CCTATTCGTG	GTCCAGATGG	GAAAATGCAT	1020
GGCAACTTGT	GTTCCATGTG	TCAAGTCTAC	TTCCAAGCAG	AAAATGAAGA	AAAGAAAAAG	1080
GCTGAAGCAC	GAGCTAGAAA	CAAAAAGAGAA	TCTGGAAAAAG	CAACCTCATA	TGCAGAGCTT	1140
TGCAATGAAT	ATCGAAAGCT	TGTGAGGAAC	GGAAAACCTG	CTTGCACCAG	AGAGAACGAT	1200
CCTATTCAGG	GCCCAGATGG	GAAAGTGCAC	GGCAACACCT	GCTCCATGTG	TGAGGTTTTT	1260
TTCCAAGCAG	AAGAAGAAGA	AAAGAAAAAG	AAGGAAGGCG	AATCAAGAAA	CAAAAGACAA	1320
TCTAAGAGTA	CAGCTTCCTT	TGAGGAGTTG	TGTAGTGAAT	ACCGCAAATC	CAGGAAAAAC	1380
GGACGGCTTT	TTTGCACCAG	AGAGAATGAC	CCCATCCAGG	GCCCAGATGG	GAAAATGCAT	1440
GGCAACACCT	GCTCCATGTG	TGAGGCCTTC	TTTCAACAAG	AAGAAAGAGC	AAGAGCAAAG	1500
GCTAAAAGAG	AAGCTGCAAA	GGAAATCTGC	AGTGAATTTT	GGGACCAAGT	GAGGAATGGA	1560
ACACTTATAT	GCACCAGGGA	GCATAATCCT	GTCCGTGGAC	CAGATGGCAA	AATGCATGGA	1620
AACAAGTGTG	CCATGTGTGC	CAGTGTGTTC	AACTTGAAG	AAGAAGAGAA	GAAAAATGAT	1680
AAAGAAGAAA	AAGGGAAAGT	TGAGGCTGAA	AAAGTTAAGA	GAGAAGCAGT	TCAGGAGCTG	1740
TGCAGTGAAT	ATCGTCATTA	TGTGAGGAAT	GGACGACTCC	CCTGTACCAG	AGAGAATGAT	1800
CCTATTGAGG	GTCTAGATGG	GAAAATCCAC	GGCAACACCT	GCTCCATGTG	TGAAGCCTTC	1860
TTCCAGCAAG	AAGCAAAAGA	AAAAGAAAGA	GCTGAACCCA	GAGCAAAAGT	CAAAAGAGAA	1920
GCTGAAAAGG	AGACATGCGA	TGAATTTTCGG	AGACTTTTGC	AAAATGGAAA	ACTTTTCTGC	1980
ACAAGAGAAA	ATGATCCTGT	GCGTGGCCCA	GATGGCAAGA	CCCATGGCAA	CAAGTGTGCC	2040
ATGTGTAAGG	CAGTCTTCCA	GAAAGAAAAT	GAGGAAAGAA	AGAGGAAAGA	AGAGGAAGAT	2100
CAGAGAAATG	CTGCAGGACA	TGGTTCCAGT	GGTGGTGGAG	GAGGAAACAC	TCAGGACGAA	2160
TGTGCTGAGT	ATCGGGAACA	AATGAAAAAT	GGAAGACTCA	GCTGTACTCG	GGAGAGTGAT	2220
CCTGTACGTG	ATGCTGATGG	CAAATCGTAC	AACAATCAGT	GTACCATGTG	TAAAGCAAAA	2280
TTGGAAAGAG	AAGCAGAGAG	AAAAAATGAG	TATTCTCGCT	CCAGATCAAA	TGGGACTGGA	2340
TCAGAATCAG	GGAAGGATAC	ATGTGATGAG	TTTAGAAGCC	AAATGAAAAA	TGGAAAACCT	2400
ATCTGCACTC	GAGAAAAGTGA	CCCTGTCCGG	GGTCCAGATG	GCAAGACACA	TGGTAATAAG	2460
TGTACTATGT	GTAAGGAAAA	ACTGGAAAGG	GAAGCAGCTG	AAAAAAAAAG	AAAGAGGATG	2520
AAGACAGGAG	CAATACAGGA	GAAAGGAGCA	ATACAGGAGA	AAGGAGCAAT	GACAAAGAGG	2580
ATCTGTGTCG	TGAATTTTCGA	AGCATGCAGA	GAAATGGAAA	GCTTATCTGC	ACCAGAGAAA	2640
ATAACCCTGT	TCGAGGCCCA	TATGGCAAGA	TGCACATCAA	TAAATGTGCT	ATGTGTCAGA	2700

GCATCTTTGA TCGAGAAGCT AATGAAAGAA AAAAGAAAGA TGAAGAGAAA TCAAGTAGCA	2760
AGCCCTCAAA TAATGCAAAG GATGAGTGCA GTGAATTTTCG AAACATATATA AGGAACAATG	2820
AACTCATCTG CCCTAGAGAG AATGACCCAG TGCACGGTGC TGATGGAAAG TTCTATACAA	2880
ACAAGTGCTA CATGTGCAGA GCTGTCTTTC TAACAGAAGC TTTGGAAAGG GCAAAGCTTC	2940
AAGAAAAACC ATCCCATGTT AGAGCTTCTC AAGAGGAAGA CAGCCCAGAC TCTTTCAGTT	3000
CTCTGGATTC TGAGATGTGC AAAGACTACC GAGTATTGCC CAGGATAGGC TATCTTTGTC	3060
CAAAGGATTT AAAGCCTGTC TGTGGTGACG ATGGCCAAAC CTACAACAAT CCTTGCATGC	3120
TCTGTCATGA AAACCTGATA CGCCAAACAA ATACACACAT CCGCAGTACA GGGAAAGTGTG	3180
AGGAGAGCAG CACCCCAGGA ACCACCGCAG CCAGCATGCC CCCGTTTGAC GAATGACAGG	3240
AAGATTGTTG AAAGCCATGA GGGAAAAAAT AAACCCAGT TTTGAATCAC CTACCTTCAC	3300
CATCTGTATA TACAAAGAAT TTTTCGGAGC TTGTTTTATT TGCTATAGAA AACAATACAG	3360
AGCTTTTGGG AATGGAATCA CTGATTTTCA GTCTTTTCCA TTTCTTTCCT CCTAGAATCT	3420
GTGATCTGAG GGTATAAAGA CATTTCACC AAGTTTGAGC CCTCAAAATG TCCTGATTAC	3480
AATGCTGTCT GTCCAATGC CTGTTCAATA AAAGTAACT CAGCAGAAAA A	3531

(2) ANGABEN ZU SEQ ID NO: 7:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

His	Glu	Phe	Gln	Ala	Phe	Met	Lys	Asn	Gly	Lys	Leu	Phe
1				5					10			

(2) ANGABEN ZU SEQ ID NO: 8:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

Ser	Glu	Tyr	Arg	Lys	Ser	Arg	Lys	Asn	Gly	Arg	Leu	Phe
1				5					10			

(2) ANGABEN ZU SEQ ID NO: 9:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 13 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

Asp Asp Phe Lys Lys Gly Glu Arg Asp Gly Asp Phe Ile
 1 5 10

(2) ANGABEN ZU SEQ ID NO: 10:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 13 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

Ser Glu Phe Arg Asp Gln Val Arg Asn Gly Thr Leu Ile
 1 5 10

(2) ANGABEN ZU SEQ ID NO: 11:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 13 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

Ser Ala Phe Arg Pro Phe Val Arg Asn Gly Arg Leu Gly
 1 5 10

(2) ANGABEN ZU SEQ ID NO: 12:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 13 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

11.04.02.99

- 19 -

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

Ser Glu Tyr Arg His Tyr Val Arg Asn Gly Arg Leu Pro
1 5 10

(2) ANGABEN ZU SEQ ID NO: 13:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

Lys Glu Tyr Glu Lys Gln Val Arg Asn Gly Arg Leu Phe
1 5 10

(2) ANGABEN ZU SEQ ID NO: 14:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

Asp Glu Phe Arg Arg Leu Leu Gln Asn Gly Lys Leu Phe
1 5 10

(2) ANGABEN ZU SEQ ID NO: 15:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

Ser Gln Tyr Gln Asn Gln Ala Lys Asn Gly Ile Leu Phe
1 5 10

(2) ANGABEN ZU SEQ ID NO: 16:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

Ala Glu Tyr Arg Glu Gln Met Lys Asn Gly Arg Leu Ser
1 5 10

(2) ANGABEN ZU SEQ ID NO: 17:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

Asn Glu Tyr Arg Lys Leu Val Arg Asn Gly Lys Leu Ala
1 5 10

(2) ANGABEN ZU SEQ ID NO: 18:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

Asp Glu Phe Arg Ser Gln Met Lys Asn Gly Lys Leu Ile
1 5 10

(2) ANGABEN ZU SEQ ID NO: 19:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 18 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

Pro Gln Asp Lys Lys Phe Phe Gln Ser Leu Asp Gly Ile Met Phe Ile
1 5 10 15

Asn Lys

(2) ANGABEN ZU SEQ ID NO: 20:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

Thr	Arg	Glu	Asn	Asp	Pro	Ile	Gln	Gly	Pro	Asp	Gly	Lys	Met	His	Gly
1				5					10					15	

Asn Thr

(2) ANGABEN ZU SEQ ID NO: 21:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

Thr	Arg	Glu	Asn	Asp	Pro	Val	Leu	Gly	Pro	Asp	Gly	Lys	Thr	His	Gly
1				5					10					15	

Asn Lys

(2) ANGABEN ZU SEQ ID NO: 22:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

Thr	Arg	Glu	His	Asn	Pro	Val	Arg	Gly	Pro	Asp	Gly	Lys	Met	His	Gly
1				5					10					15	

Asn Lys

(2) ANGABEN ZU SEQ ID NO: 23:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 18 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

Thr Arg Glu Ser Asp Pro Val Arg Gly Pro Asp Gly Arg Met His Gly
 1 5 10 15

Asn Lys

(2) ANGABEN ZU SEQ ID NO: 24:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 18 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

Thr Arg Glu Asn Asp Pro Ile Glu Gly Leu Asp Gly Lys Ile His Gly
 1 5 10 15

Asn Thr

(2) ANGABEN ZU SEQ ID NO: 25:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 18 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

Thr Arg Glu Asn Asp Pro Ile Arg Gly Pro Asp Gly Lys Met His Gly
 1 5 10 15

Asn Leu

(2) ANGABEN ZU SEQ ID NO: 26:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

Thr	Arg	Glu	Asn	Asp	Pro	Val	Arg	Gly	Pro	Asp	Gly	Lys	Thr	His	Gly
1				5					10					15	

Asn Lys

(2) ANGABEN ZU SEQ ID NO: 27:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

Thr	Arg	Glu	Asn	Asp	Pro	Ile	Gln	Gly	Pro	Asp	Gly	Lys	Val	His	Gly
1				5					10					15	

Asn Thr

(2) ANGABEN ZU SEQ ID NO: 28:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

Thr	Arg	Glu	Ser	Asp	Pro	Val	Arg	Asp	Ala	Asp	Gly	Lys	Ser	Tyr	Asn
1				5					10					15	

Asn Gln

(2) ANGABEN ZU SEQ ID NO: 29:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 18 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

Thr Arg Glu Ser Asp Pro Val Arg Gly Pro Asp Gly Lys Thr His Gly
 1 5 10 15
 Asn Lys

(2) ANGABEN ZU SEQ ID NO: 30:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 2 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

Ala Thr
 1

(2) ANGABEN ZU SEQ ID NO: 31:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 2 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

Ala Leu
 1

(2) ANGABEN ZU SEQ ID NO: 32:

- (i) SEQUENZKENNZEICHEN:
 - (A) LÄNGE: 2 Aminosäuren
 - (B) ART: Aminosäure
 - (C) STRANGFORM: Einzelstrang
 - (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

Ala Met
1

(2) ANGABEN ZU SEQ ID NO: 33:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 2 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

Ser Met
1

(2) ANGABEN ZU SEQ ID NO: 34:

- (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 2 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: nicht bekannt

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

Thr Met
1

A n s p r ü c h e

1. Serin-Proteinase-Inhibitor, dadurch gekennzeichnet, daß der Serin-Proteinase-Inhibitor eine Domäne mit vier Cysteinen aufweist und sich zwischen einem ersten und einem zweiten Cystein eine Sequenz von 0 bis 20 Aminosäuren befindet oder der Serin-Proteinase-Inhibitor eine Domäne mit sechs Cysteinen aufweist und sich zwischen einem ersten und zweiten Cystein eine Sequenz von 7 bis 20 Aminosäuren befindet.
2. Serin-Proteinase-Inhibitor gemäß Anspruch 1, dadurch gekennzeichnet, daß sich zwischen dem ersten und dem zweiten Cystein der Domäne eine Sequenz von 13 Aminosäuren befindet.
3. Serin-Proteinase-Inhibitor gemäß Anspruch 1 und/oder 2, dadurch gekennzeichnet, daß sich zwischen dem zweiten und dem dritten Cystein der Domäne eine Sequenz von 18 Aminosäuren befindet.
4. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß sich zwischen dem dritten und vierten Cystein der Domäne eine Sequenz von 2 Aminosäuren befindet.
5. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß die Sequenz der Domäne zwischen dem ersten und dem zweiten Cystein ausgewählt wird aus

HEFQAFMKNGKLF,
DDFKKGERDGDFI,
SAFRPFVRNGRLG,
KEYEKQVRNGRLF,
SQYQNQAKNGILF,

SEYRKSRRKNGRLF,
SEFRDQVRNGTLI,
SEYRHYVRNGRLP,
DEFRRLLQNGKLF,
AEYREQMKNGRLS oder

NEYRKLVRNGKLA,

DEFRSQMKNGLI.

6. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die Sequenz zwischen dem zweiten und dem dritten Cystein der Domäne ausgewählt wird aus

PQDKKFFQSLDGIMFINK,	TRENDPIQGPDPGKMHGNT,
TRENDPVLGPDGKTHGNK,	TREHNPVRGPDGKMHGNT,
TRESDPVRGPDGRMHGNT,	TRENDPIEGLDGKIHGNT,
TRENDPIRGPDGKMHGNT,	TRENDPVRGPDGKTHGNK,
TRENDPIQGPDPGKVHGNT,	TRESDPVRDADGKSYNNQ oder
TRESDPVRGPDGKTHGNK.	

7. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß die Sequenz zwischen dem dritten und dem vierten Cystein der Domäne ausgewählt wird aus

AT, AL, AM, SM oder TM.

8. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 7, mit einer der Formeln

R_1 -C-HEFQAFMKNGLF-C-PQDKKFFQSLDGIMFINK-C-AT-C- R_2
 R_1 -C-DDFKKGERDGDGI-C-PDYVEAVCGTDGKTYDNR-C-AL-C- R_2
 R_1 -C-SAFRPFVRNGRLG-C-TRENDPVLGPDGKTHGNK-C-AM-C- R_2
 R_1 -C-KEYEKQVRNGRLF-C-TRESDPVRGPDGRMHGNT-C-AL-C- R_2
 R_1 -C-SQYQNQAKNGILF-C-TRENDPIRGPDGKMHGNT-C-SM-C- R_2
 R_1 -C-NEYRKLVRNGKLA-C-TRENDPIQGPDPGKVHGNT-C-SM-C- R_2
 R_1 -C-SEYRKSRRNGRLF-C-TRENDPIQGPDPGKMHGNT-C-SM-C- R_2
 R_1 -C-SEFRDQVRNGTLI-C-TREHNPVRGPDGKMHGNT-C-AM-C- R_2
 R_1 -C-SEYRHYVRNGRLP-C-TRENDPIEGLDGKIHGNT-C-SM-C- R_2
 R_1 -C-DEFRRLLQNGKLF-C-TRENDPVRGPDGKTHGNK-C-AM-C- R_2
 R_1 -C-AEYREQMKNGLS-C-TRESDPVRDADGKSYNNQ-C-TM-C- R_2
 R_1 -C-DEFRSQMKNGLI-C-TRESDPVRGPDGKTHGNK-C-TM-C- R_2 ,

worin R_1 NH_2 , eine Aminosäure oder ein Peptid mit bis zu 1.000 Aminosäuren ist und R_2 $COOH$, $CONH_2$, eine Aminosäure oder ein Peptid mit bis zu 1.000 Aminosäuren ist.

9. Serin-Proteinase-Inhibitor gemäß mindestens einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß
 - er zwischen dem ersten und dem vierten Cystein und/oder dem zweiten und dem dritten Cystein eine Disulfidbrücke aufweist oder
 - er zwischen dem ersten und dem fünften und/oder dem zweiten und dem vierten und/oder dem dritten und dem sechsten Cystein eine Disulfidbrücke aufweist.
10. Serin-Proteinase-Inhibitor gemäß mindestens einem der Ansprüche 1 bis 9, dadurch gekennzeichnet, daß es sich um ein Fragment von VAKTI-1 (Seq. ID Nr. 1) oder VAKTI-2 (Seq. ID Nr. 2) handelt.
11. Serin-Proteinase-Inhibitor gemäß Anspruch 10, dadurch gekennzeichnet, daß es sich um HF 6479 (Seq. ID Nr. 3) oder HF 7665 (Seq. ID Nr. 4) handelt.
12. Nucleinsäure, kodierend für einen Serin-Proteinase-Inhibitor gemäß mindestens einem der Ansprüche 1 bis 11.
13. Arzneimittel enthaltend mindestens einen Serin-Proteinase-Inhibitor gemäß mindestens einem der Ansprüche 1 bis 11 und/oder eine Nucleinsäure gemäß Anspruch 12 gegebenenfalls mit pharmazeutischen Trägerstoffen.
14. Arzneimittel nach Anspruch 13, enthaltend 0,01 bis 1.000 mg/kg Körpergewicht des Serin-Proteinase-Inhibitors nach mindestens einem der Ansprüche 1 bis 11 und/oder der Nucleinsäure gemäß Anspruch 12.

15. Verwendung des Serin-Proteinase-Inhibitors nach mindestens einem der Ansprüche 1 bis 11 zur Herstellung eines Arzneimittels zur Behandlung von akuten oder chronischen Cervixentzündungen, Entzündungen der Bartholinschen Drüsen und anderer vaginaler Bereiche, Tonsillitis, Pharyngitis und Laryngitis, mit exzessiver Schleimbildung verbundener akut oder chronisch entzündlicher Prozesse und sich daraus ergebender akuter Notsituationen, postoperativer Blutungen aufgrund Hyperfibrinolyse sowie zur Prophylaxe der Lungenemphysembildung bei α_1 -Proteinase-Inhibitor-mangel.
16. Verwendung der Nucleinsäuren gemäß Anspruch 12 zur Herstellung eines Arzneimittels zum Einsatz in der Gentherapie zur Heilung und Prophylaxe von Erkrankungen gemäß Anspruch 12.
17. Antikörper oder Antikörper-Fragmente gegen Epitope der Verbindungen nach einem der Ansprüche 1 bis 11.
18. Poly- oder Oligonucleotide, die mit Bereichen der cDNA oder entsprechender RNA unter stringenten Bedingungen hybridisieren und gegebenenfalls die Expression codierender Bereiche der für die Verbindungen nach Anspruch 1 bis 11 codierenden Gene verhindern (Antisense Verbindungen).
19. Diagnostikum enthaltend mindestens eine der Verbindungen gemäß Anspruch 17 oder 18.
20. Arzneimittel enthaltend mindestens eine der in den Ansprüchen 17 und/oder 18 genannten Verbindungen in therapeutisch wirksamen Mengen.
21. Verwendung der Verbindungen gemäß Anspruch 17 und/oder 18 zur Herstellung eines Arzneimittels zur Behandlung von Erkrankungen, die mit einer zu hohen Expression der

Verbindungen gemäß mindestens einem der Ansprüche 1 bis 11 oder zu hohen Aktivität der für die Verbindungen gemäß Ansprüchen 1 bis 11 codierenden Bereichen verbunden sind.

22. DNA kodierend für die in den Ansprüchen 1 bis 11 genannten Verbindungen und/oder RNA, die in der Transcription oder Translation der in den Ansprüchen 1 bis 11 genannten Verbindungen beteiligt ist.
23. DNA gemäß Anspruch 22 mit der Seq. ID Nr. 5 oder Seq. ID Nr. 6.

Zusammenfassung

Serin-Proteinase-Inhibitor, dadurch gekennzeichnet, daß der Serin-Proteinase-Inhibitor eine Domäne mit vier Cysteinen aufweist und sich zwischen einem ersten und einem zweiten Cystein eine Sequenz von 0 bis 20 Aminosäuren befindet oder der Serin-Proteinase-Inhibitor eine Domäne mit sechs Cysteinen aufweist und sich zwischen einem ersten und zweiten Cystein eine Sequenz von 7 bis 20 Aminosäuren befindet.

Figur 1

VAKTI-1 cDNA und Translation in Aminosäuresequenz

Frame 2

ATG CAT GGA GTG GAC CTG TAG GCG ACT TGC ATC GTC TTC AAC ATG AAG ATA GCC
 10 19 28 37 46 55
 M K I A
 |—MEMC-1—>
 |—>HF6479
 T V S V L L P L A L C L I Q D A A S K N
 ACA GTG TCA GTG CTT CTG CCC TTG GCT CTT TGC CTC ATA CAA GAT GCT GCC AGT AAG AAT
 64 73 82 91 100 109
 —MEMC-1—> —CHEF-1—>
 E D Q E M C H E F Q A F M K N G K L F C
 GAA GAT CAG GAA ATG TGC CAT GAA TTT CAG GCA TTT ATG AAA AAT GGA AAA CTG TTC TGT
 124 133 142 151 160 169
 <—CHEF-14—>
 <—CHEF-11—>
 P Q D K K F F Q S L D G I M F I N K C A
 CCC CAG GAT AAG AAA TTT TTT CAA AGT CTT GAT GGA ATA ATG TTC ATC AAT AAA TGT GCC
 184 193 202 211 220 229
 —CHEF-2—> HF6479 <—|
 C K M I L E K E A K S Q K R A R H L A
 ACG TGC AAA ATG ATA CTG GAA AAA GAA GCA AAA TCA CAG AAG AGG GCC AGG CAT TTA GCA
 244 253 262 271 280 289
 R A P K A T A P T E L N C D D F K K G E
 AGA GCT CCC AAG GCT ACT GCC CCA ACA GAG CTG AAT TGT GAT GAT TTT AAA AAA GGA GAA
 304 313 322 331 340 349
 R D G D F I C P D Y Y E A V C G T D G K
 AGA GAT GGG GAT TTT ATC TGT CCT GAT TAT TAT GAA GCT GTT TGT GGC ACA GAT GGG AAA
 364 373 382 391 400 409
 T Y D N R C A L C A E N A K T G S Q I G
 ACA TAT GAC AAC AGA TGT GCA CTG TGT GCT GAG AAT GCG AAA ACC GGG TCC CAA ATT GGT
 424 433 442 451 460 469
 V K S E G E C K S S N P E Q V R S I V S
 GTA AAA AGT GAA GGG GAA TGT AAG AGC AGT AAT CCA GAG CAG GTG AGG TCA ATT GTC AGC
 484 493 502 511 520 529
 M G N T G R L T S N S K STOP
 ATG GGA AAT ACT GGG AGG CTA ACT TCA AAT AGT AAG TAG GTG CTG TCC TCT TCC TTC
 544 553 562 571 580 589
 TTA GGT GGG AGC CTT GGA AGG AAT TAA TTC TTG CTT TAT GTG AAA TGG AAT ACC CAG TTA
 604 613 622 631 640 649
 CTG CCC ACT AAT ATG AAA AAG CTA ATT ATA GTC TCT GAA ACT GGA TCA GAT TAC TTT GGT
 664 673 682 691 700 709
 GGT TAA GAT CTT TCA ATC TAT TGC TGC TTT GTA T
 724 733 742 749

Figur 2

VAKTI-2 cDNA und Translation in Aminosäuresequenz

Frame 2

ATG CAT GGA GTG GAC CTG TAG GCG ACT TGC ATC GTC TTC AAC ATG AAG ATA GCC
10 19 28 37 46 55

T V S V L L P L A L C L I Q D A A S K N
ACA GTG TCA GTG CTT CTG CCC TTG GCT CTT TGC CTC ATA CAA GAT GCT GCC AGT AAG AAT
64 73 82 91 100 109

Repeat 1

E D Q E M C H E F Q A F M K N G K L F C
GAA GAT CAG GAA ATG TGC CAT GAA TTT CAG GCA TTT ATG AAA AAT GGA AAA CTG TTC TGT
124 133 142 151 160 169
P Q D K K F F Q S L D G I M F I N K C A
CCC CAG GAT AAG AAA TTT TTT CAA AGT CTT GAT GGA ATA ATG TTC ATC AAT AAA TGT GCC
184 193 202 211 220 229

HF 6479 <—|

T C K M I L E K E A K S Q I K R A R H L A
TGC AAA ATG ATA CTG GAA AAA GAA GCA AAA TCA CAG AAG AGG GCC AGG CAT TTA GCA
244 253 262 271 280 289

Typische Kazal-Domäne

R A P K A T A P T E L N C D D F K K G E
AGA GCT CCC AAG GCT ACT GCC CCA ACA GAG CTG AAT TGT GAT GAT TTT AAA AAA GGA GAA
304 313 322 331 340 349
R D G D F I C P D Y Y E A V C G T D G K
AGA GAT GGG GAT TTT ATC TGT CCT GAT TAT TAT GAA GCT GTT TGT GGC ACA GAT GGG AAA
364 373 382 391 400 409
T Y D N R C A L C A E N A K T G S Q I G
ACA TAT GAC AAC AGA TGT GCA CTG TGT GCT GAG AAT GCG AAA ACC GGG TCC CAA ATT GGT
424 433 442 451 460 469

Repeat 2

V K S E G E C K S S N P E Q D V C S A F
GTA AAA AGT GAA GGG GAA TGT AAG AGC AGT AAT CCA GAG CAG GAT GTA TGC AGT GCT TTT
484 493 502 511 520 529
R P F V R N G R L G C T R E N D P V L G
CCC TTT GTT AGA AAT GGA AGA CTT GGA TGC ACA AGG GAA AAT GAT CCT GTT CTT GGT
544 553 562 571 580 589
P D G K T H G N K C A M C A E L F L K E
CCT GAT GGG AAG ACG CAT GGC AAT AAG TGT GCA ATG TGT GCT GAG CTG TTT TTA AAA GAA
604 613 622 631 640 649
A E N A K R E G E T R I R R N A E K D F
GCT GAA AAT GCC AAG CGA GAG GGT GAA ACT AGA ATT CGA CGA AAT GCT GAA AAG GAT TTT
664 673 682 691 700 709

Repeat 3

C K E Y E K Q V R N G R L F C T R E S D
TGC AAG GAA TAT GAA AAA CAA GTG AGA AAT GGA AGG CTT TTT TGT ACA CGG GAG AGT GAT
724 733 742 751 760 769
P V R G P D G R M H G N K C A L C A E I
CCA GTC CGT GGC CCT GAC GGC AGG ATG CAT GGC AAC AAA TGT GCC CTG TGT GCT GAA ATT
784 793 802 811 820 829
F K R R F S E E N S K T D Q N L G K A E
TTC AAG CGG CGT TTT TCA GAG GAA AAC AGT AAA ACA GAT CAA AAT TTG GGA AAA GCT GAA
844 853 862 871 880 889

Repeat 4

E K T K V K R E I V K L C S Q Y Q N Q A
 GAA AAA ACT AAA GTT AAA AGA GAA ATT GTG AAA CTC TGC AGT CAA TAT CAA AAT CAG GCA
 904 913 922 931 940 949

 K N G I L F C T R E N D P I R G P D G K
 AAG AAT GGA ATA CTT TTC TGT ACC AGA GAA AAT GAC CCT ATT CGT GGT CCA GAT GGG AAA
 964 973 982 991 1000 1009

 M H G N L C S M C Q V Y F Q A E N E E K
 ATG CAT GGC AAC TTG TGT TCC ATG TGT CAA GTC TAC TTC CAA GCA GAA AAT GAA GAA AAG
 1024 1033 1042 1051 1060 1069

|——> HF 7665

K K A E A R A R N K R I E S G K A T S Y A
 AAA AAG GCT GAA GCA CGA GCT AGA AAC AAA AGA GAA TCT GGA AAA GCA ACC TCA TAT GCA
 1084 1093 1102 1111 1120 1129

Repeat 5

 E L C N E Y R K L V R N G K L A C T R E
 GAG CTT TGC AAT GAA TAT CGA AAG CTT GTG AGG AAC GGA AAA CTT GCT TGC ACC AGA GAG
 1144 1153 1162 1171 1180 1189

 D P I Q G P D G K V H G N T C S M C E
 C GAT CCT ATT CAG GGC CCA GAT GGG AAA GTG CAC GGC AAC ACC TGC TCC ATG TGT GAG
 1204 1213 1222 1231 1240 1249

HF 7665 <——|

V F F Q A E E E E K K K K E G E S R N K
 GTT TTT TTC CAA GCA GAA GAA GAA GAA AAG AAA AAG AAG GAA GGC GAA TCA AGA AAC AAA
 1264 1273 1282 1291 1300 1309

Repeat 6

 R Q S K S T A S F E E L C S E Y R K S R
 AGA CAA TCT AAG AGT ACA GCT TCC TTT GAG GAG TTG TGT AGT GAA TAC CGC AAA TCC AGG
 1324 1333 1342 1351 1360 1369

 K N G R L F C T R E N D P I Q G P D G K
 AAA AAC GGA CGG CTT TTT TGC ACC AGA GAG AAT GAC CCC ATC CAG GGC CCA GAT GGG AAA
 1384 1393 1402 1411 1420 1429

 M H G N T C S M C E A F F Q Q E E R A R
 ATG CAT GGC AAC ACC TGC TCC ATG TGT GAG GCC TTC TTT CAA CAA GAA GAA AGA GCA AGA
 1444 1453 1462 1471 1480 1489

Repeat 7

 K A K R E A A K E I C S E F R D Q V R
 GCA AAG GCT AAA AGA GAA GCT GCA AAG GAA ATC TGC AGT GAA TTT CGG GAC CAA GTG AGG
 1504 1513 1522 1531 1540 1549

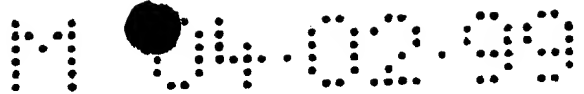
 N G T L I C T R E H N P V R G P D G K M
 AAT GGA ACA CTT ATA TGC ACC AGG GAG CAT AAT CCT GTC CGT GGA CCA GAT GGC AAA ATG
 1564 1573 1582 1591 1600 1609

 H G N K C A M C A S V F K L E E E E K K
 CAT GGA AAC AAG TGT GCC ATG TGT GCC AGT GTG TTC AAA CTT GAA GAA GAA GAG AAG AAA
 1624 1633 1642 1651 1660 1669

 N D K E E K G K V E A E K V K R E A V Q
 AAT GAT AAA GAA GAA AAA GGG AAA GTT GAG GCT GAA AAA GTT AAG AGA GAA GCA GTT CAG
 1684 1693 1702 1711 1720 1729

Repeat 8

 E L C S E Y R H Y V R N G R L P C T R E
 GAG CTG TGC AGT GAA TAT CGT CAT TAT GTG AGG AAT GGA CGA CTC CCC TGT ACC AGA GAG
 1744 1753 1762 1771 1780 1789



N D P I E G L D G K I H G N T C S M C E
 AAT GAT CCT ATT GAG GGT CTA GAT GGG AAA ATC CAC GGC AAC ACC TGC TCC ATG TGT GAA
 1804 1813 1822 1831 1840 1849

A F F Q Q E A K E K E R A E P R A K V K
 GCC TTC TTC CAG CAA GAA GCA AAA GAA AAA GAA AGA GCT GAA CCC AGA GCA AAA GTC AAA
 1864 1873 1882 1891 1900 1909

Repeat 9

R E A E K E T C D E F R R L L Q N G K L
 AGA GAA GCT GAA AAG GAG ACA TGC GAT GAA TTT CGG AGA CTT TTG CAA AAT GGA AAA CTT
 1924 1933 1942 1951 1960 1969

F C T R E N D P V R G P D G K T H G N K
 TTC TGC ACA AGA GAA AAT GAT CCT GTG CGT GGC CCA GAT GGC AAG ACC CAT GGC AAC AAG
 1984 1993 2002 2011 2020 2029

C A M C K A V F Q K E N E E R K R K E E
 TGT GCC ATG TGT AAG GCA GTC TTC CAG AAA GAA AAT GAG GAA AGA AAG AGG AAA GAA GAG
 2044 2053 2062 2071 2080 2089

D Q R N A A G H G S S G G G G G N T Q
 AAT GAT CAG AGA AAT GCT GCA GGA CAT GGT TCC AGT GGT GGT GGA GGA GGA AAC ACT CAG
 2104 2113 2122 2131 2140 2149

Repeat 10

D E C A E Y R E Q M K N G R L S C T R E
 GAC GAA TGT GCT GAG TAT CGG GAA CAA ATG AAA AAT GGA AGA CTC AGC TGT ACT CGG GAG
 2164 2173 2182 2191 2200 2209

S D P V R D A D G K S Y N N Q C T M C K
 AGT GAT CCT GTA CGT GAT GCT GAT GGC AAA TCG TAC AAC AAT CAG TGT ACC ATG TGT AAA
 2224 2233 2242 2251 2260 2269

A K L E R E A E R K N E Y S R S R S N G
 GCA AAA TTG GAA AGA GAA GCA GAG AGA AAA AAT GAG TAT TCT CGC TCC AGA TCA AAT GGG
 2284 2293 2302 2311 2320 2329

Repeat 11

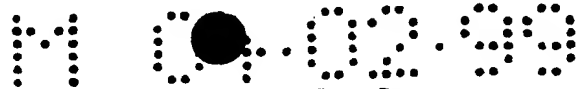
G S E S G K D T C D E F R S Q M K N G
 GGA TCA GAA TCA GGG AAG GAT ACA TGT GAT GAG TTT AGA AGC CAA ATG AAA AAT GGA
 2344 2353 2362 2371 2380 2389

K L I C T R E S D P V R G P D G K T H G
 AAA CTT ATC TGC ACT CGA GAA AGT GAC CCT GTC CGG GGT CCA GAT GGC AAG ACA CAT GGT
 2404 2413 2422 2431 2440 2449

N K C T M C K E K L E R E A A E K K R K
 AAT AAG TGT ACT ATG TGT AAG GAA AAA CTG GAA AGG GAA GCA GCT GAA AAA AAA AGA AAG
 2464 2473 2482 2491 2500 2509

R M K T G A I Q E K G A I Q E K G A M T
 AGG ATG AAG ACA GGA GCA ATA CAG GAG AAA GGA GCA ATA CAG GAG AAA GGA GCA ATG ACA
 2524 2533 2542 2551 2560 2569

K R I C V V N F E A C R E M E S L S A P
 AAG AGG ATC TGT GTC GTG AAT TTC GAA GCA TGC AGA GAA ATG GAA AGC TTA TCT GCA CCA
 2584 2593 2602 2611 2620 2629



E K I T L F E A H M A R C T S I N V L C
 GAG AAA ATA ACC CTG TTC GAG GCC CAT ATG GCA AGA TGC ACA TCA ATA AAT GTG CTA TGT
 2644 2653 2662 2671 2680 2689

V R A S L I E K L M K E K R K M K R N Q
 GTC AGA GCA TCT TTG ATC GAG AAG CTA ATG AAA GAA AAA AGA AAG ATG AAG AGA AAT CAA
 2704 2713 2722 2731 2740 2749

V A S P Q I M Q R M S A V N F E T I STOP
 GTA GCA AGC CCT CAA ATA ATG CAA AGG ATG AGT GCA GTG AAT TTC GAA ACT ATA TAA GGA
 2764 2773 2782 2791 2800 2809

ACA ATG AAC TCA TCT GCC CTA GAG AGA ATG ACC CAG TGC ACG GTG CTG ATG GAA AGT TCT
 2824 2833 2842 2851 2860 2869

ATA CAA ACA AGT GCT ACA TGT GCA GAG CTG TCT TTC TAA CAG AAG CTT TGG AAA GGG CAA
 2884 2893 2902 2911 2920 2929

C TTC AAG AAA AAC CAT CCC ATG TTA GAG CTT CTC AAG AGG AAG ACA GCC CAG ACT CTT
 2944 2953 2962 2971 2980 2989

TCA GTT CTC TGG ATT CTG AGA TGT GCA AAG ACT ACC GAG TAT TGC CCA GGA TAG GCT ATC
 3004 3013 3022 3031 3040 3049

TTT GTC CAA AGG ATT TAA AGC CTG TCT GTG GTG ACG ATG GCC AAA CCT ACA ACA ATC CTT
 3064 3073 3082 3091 3100 3109

GCA TGC TCT GTC ATG AAA ACC TGA TAC GCC AAA CAA ATA CAC ACA TCC GCA GTA CAG GGA
 3124 3133 3142 3151 3160 3169

AGT GTG AGG AGA GCA GCA CCC CAG GAA CCA CCG CAG CCA GCA TGC CCC CGT TTG ACG AAT
 3184 3193 3202 3211 3220 3229

AGG AAG ATT GTT GAA AGC CAT GAG GGA AAA AAT AAA CCC CAG TTT TGA ATC ACC TAC
 3244 3253 3262 3271 3280 3289

CTT CAC CAT CTG TAT ATA CAA AGA ATT TTT CGG AGC TTG TTT TAT TTG CTA TAG AAA ACA
 3304 3313 3322 3331 3340 3349

ATA CAG AGC TTT TGG GAA TGG AAT CAC TGA TTT TCA GTC TTT TCC ATT TCT TTC CTC CTA
 3364 3373 3382 3391 3400 3409

GAA TCT GTG ATC TGA GGG TAT AAA GAC ATT TCC ACC AAG TTT GAG CCC TCA AAA TGT CCT
 3424 3433 3442 3451 3460 3469

Polyadenylierungssignal

GAT TAC AAT GCT GTC TGT CCA ACT GCC TGT TCA ATA AAA GTA AAC TCA GCA GAA AAA....
 3484 3493 3502 3511 3520 3529

.....Poly(A)-Tail

Figur 3

